LA-UR-02-0014

Approved for public release;
distribution is unlimited.

Title:	FUNCTIONAL SENSITIVITY ANALYSIS FOR COMPUTER MODEL OUTPUT
Author(s):	Katherine Campbell
Submitted to:	Proceedings of the Seventh Army Conference on Statistics, Santa Fe, NM, October 22-26, 2001

Functional Sensitivity Analysis for Computer Model Output

Katherine Campbell
Statistical Sciences Group
Los Alamos National Laboratory
Los Alamos, NM 87545

Abstract

The outputs of computational models are often time series or functions of other continuous variables (space, angle, etc.) For the purposes of model sensitivity and uncertainty analysis, it makes little sense to treat individual points on these curves as scalars. Of much greater interest is the effect of model input choices and uncertainties on the overall shapes of such curves. We explore a range of methods for characterizing a set of functions generated by a series of model runs for the purposes of exploring the relationships between these function and the model inputs.

Introduction

The outputs of computational models are often time series or functions of other continuous variables (space, angle, etc.) In this paper, we propose that sensitivity analysis of such outputs be carried out by means of expansion of the functional outputs in an appropriate functional coordinate system, i.e., in terms of an appropriate set of basis functions, followed by sensitivity analysis using any standard method of the coefficients of the expansion. The only new problem, therefore, is choosing an appropriate coordinate system in which to apply the selected sensitivity analysis methods. We consider both pre-defined basis sets and data-adaptive basis sets, with their associated advantages and disadvantages. We devote only passing mention to some related, but important problems, such as increasing the interpretability of the results by appropriate preprocessing of the functional outputs (in particular, curve registration), and by enforcing some degree of smoothness when data-adaptive bases are used.

Figure 1 shows a simple made-up example. This is a set of four-parameter curves, that is, the model in this case is just a function with four parameters, *a*, *b*, *c* and *d*:

$$f(\theta) = 10 + a \exp\left(-\frac{(\theta - b)^2}{\mathrm{K}_1 a^2 + c^2} \right) + (b + d) \exp\left(\mathrm{K}_2 \, a \, \theta\right). \qquad (1)$$

We interpret these functions as output from a problem with azimuthal symmetry, say a shock wave problem or an implosion problem. The independent variable θ is a polar angle ranging from -90º to 90º. The model was run 81 times, using a complete 3^4 factorial design for the four input parameters. The 81 output curves are color coded in Figure 1 according to the value of the parameter *a*. Eq. (1) (which of course in a real example would be unknown) shows that this parameter controls the height of the central peak (and also, but less strongly, its width as well as the scale of the right-hand tail.)

In analyzing this "model output" we are typically less interested in what affects the values at, say, 45 degrees, than in questions such as: what shifts the curves up and down? left or right? What makes the central peak wider or narrower? the right-hand tail higher or lower? We could, of course, pick some appropriate functionals for answering these questions. The last, for example, we might address by examining the sensitivity of the

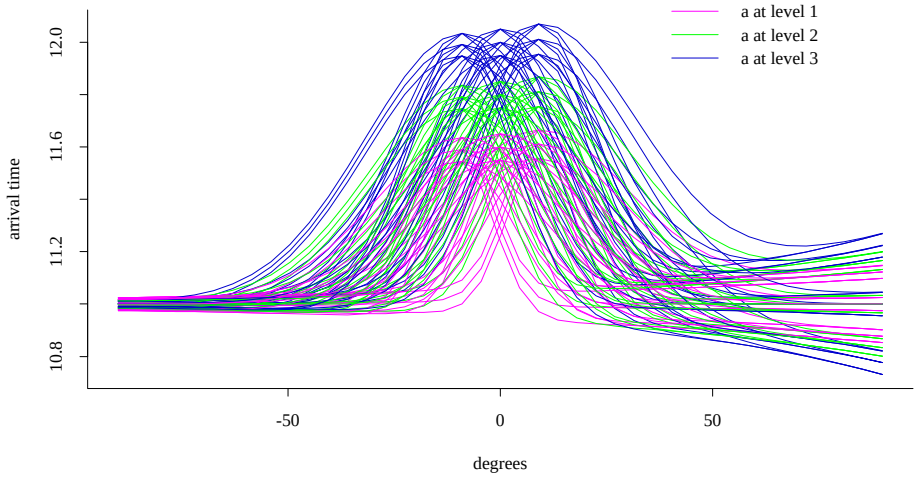

Figure 1. 81 runs of the four-parameter (4-P) model

values at 90º to the four input parameters. In order to address questions such as peak width we could devise some surrogate measurement that could be computed on each curve and then study its sensitivity to the input parameters. However, such choices are highly problem specific.

The curves in Figure 2 are outputs from 102 runs of an accelerator beam transport model, from a study that varied 18 input parameters. Twelve of these parameters describe the shape of the input beam in six dimensions (x and y position and momentum, phase and energy), while six are perturbations of the parameters of the transport line elements (x and y position and angular orientation of two quadrupole magnets). The output profiles are

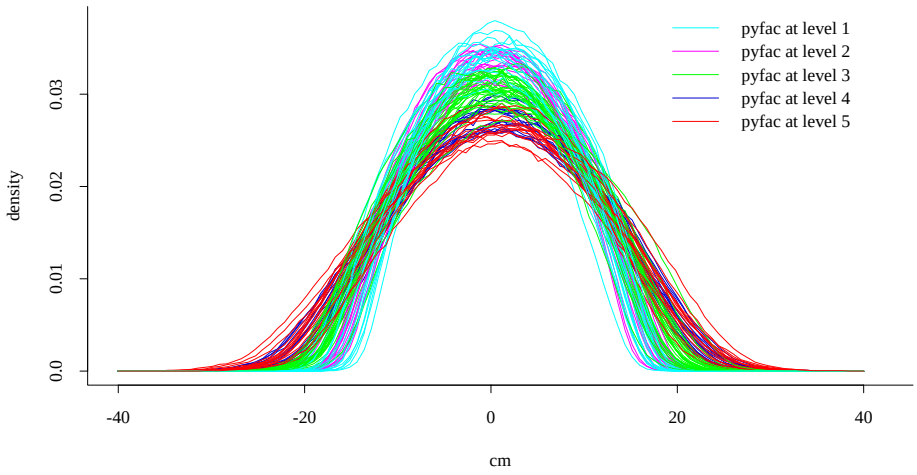

Figure 2. 102 runs of the beam transport model

compared with measurements made by a wire scanner inserted into the beam, moving from one side to the other of the beam. That is, one scan provides T=41 values along a one-dimensional projection of the beam intensity at a given point along the length of the transport line, a function of one variable which we again call θ. All the curves are normalized to have the same area underneath them.

With the benefit of hindsight, the curves have been color-coded according to the level of the input parameter which determines the width of y momentum distribution in the input beam. The principal effect of this parameter is to make the basically Gaussian shape of the intensity distribution of the output beam at the wire scanner either wider or narrower.

Transforming functional data

It might seem natural to regard functions provided on a grid of T points as T dependent variables for the purposes of sensitivity analysis. However, this is unsatisfactory for many reasons:

- The T variables are highly correlated with one another, so this natural coordinate system is inefficient for statistical methods like discriminant analysis, sensitivity analysis, or almost anything other multivariate statistical method. Results are redundant from one value of θ to another.
- The results obtained in this way are often not particularly interpretable for the underlying physical or modeling problem.
- Even though the data are the output of a computer model, the different runs may not have generated outputs at the same times or points θ. Alternatively, identical model output times may not be physically comparable because, as a function of the input parameters, the modeled process may be evolving faster in one run than another. So we may need to register the output curves (rescale time) in some physically more interpretable manner before proceeding with analysis.
- Finally, smoothness of the true curves may be a physical expectation that is not preserved by multivariate procedures in the original coordinate systems.

All of these problems can be addressed by transforming the functional output in one way or another. For sensitivity analysis, the most useful approach is expanding the output functions in terms of some basis functions (after rescaling time, if necessary) and then applying the statistical method of interest—in our case, a sensitivity analysis method—to the coefficients of that expansion. Different types of bases can be considered. There are familiar, predefined bases such as Legendre polynomials or other orthogonal polynomials, trigonometric functions, Haar functions, or wavelet bases. Adaptive basis functions include principal components, partial least squares components, and bases extracted adaptively from overcomplete dictionaries. If the columns of $\Phi_{T\times K}$ ($K \le T$) are a proposed set of basis functions, then the original functional output from N model runs, an $N \times T$ matrix Y, can be rewritten as

$$Y - \bar{Y} = H\Phi^T, \text{ i.e., } y_i(t) - \bar{y}(t) = \sum_{k=1}^{K} h_{ik}\varphi_k(t) \text{ for } 1 \le i \le N, \qquad (2)$$

where the mean function $\bar{y}(t)$ is computed as the mean of the $y_i(t)$ for each t.

Most standard basis systems are orthonormal. For example, the Legendre polynomials are orthonormal with respect to Lebesgue measure on [-1,1]. But the Legendre polynomials in sin(t), which are used in the examples below, are not orthonormal with

respect to ordinary Lebesgue measure $d\theta$, but only with respect to a weighted measure $\cos\theta d\theta$. Adaptive bases functions may be orthonormal by construction, or not. Orthonormality of the basis functions is a nice property, since then the total variance is naturally partitioned among the variances of the coefficients:

$$\sum_{i=1}^{N}\|y_i\|^2 = \sum_{i=1}^{N}\left(\sum_{t=1}^{T} y_i(t)^2\right) \sim \sum_{i=1}^{N}\left(\sum_{k=1}^{K} h_{ik}^2\right) = \sum_{i=1}^{N}\|h_i\|^2 . \qquad (3)$$

(Usually the basis functions are ordered so that the first few capture most of the total variance.) However, even when the basis functions are not orthonormal the total variance captured by the expansion in terms of the first k ($k \le K$) basis functions can be computed, and orthonormality may be less important than some other features when it comes to sensitivity analysis.

Legendre analysis for 4-P example

Since the first example is being interpreted as a set of functions of angles from $-\pi/2$ to $\pi/2$, the Legendre expansion in sin(t) is a natural choice among standard expansions. Figure 3 shows how the coefficients $\{h_{ik}\}$ of the expansions of the 81 functional outputs depend on the parameters, for k=1, 2, ..., 6. The Legendre polynomials are alternately symmetric and anti-symmetric around zero, as shown in the top row of Figure 3. The first k polynomials define a k-dimensional subspace of the 41-dimensional space in which the output functions are vectors. The percentages at the top show how much of the total variance in the original family of functions lies in this subspace for k up to 6. Note for future reference that the six-dimensional subspace defined by the first six polynomials still includes less than 90% of the total.

In the second row, the Legendre polynomials are interpreted as perturbations of the overall mean of the 81 output functions. The mean function is shown in blue, the mean plus a multiple of the Legendre polynomial in green, and the mean minus the same multiple in magenta.

The remaining rows contain box plots showing dependencies of the coefficients on the four parameters. Of course, we are not proposing sensitivity analysis by inspection as a serious method for sensitivity analysis, but SA methodology is not the main goal of this paper. The figures are intended to suggest what more formal sensitivity analysis would indicate. For the beam transport example, the displayed subset of five of the 18 input parameters was selected based on the results the partial correlation coefficient (cf. McKay, 1997, for example) on the PLS components.

For the 4-P example, variability in the coefficients of the Legendre polynomials of even order is controlled largely by *a*, although *c* and *d* influence the constant, zero-order term. The odd orders are controlled mostly by *b* with some influence of *d* on the first-order term.

Legendre polynomials and other standard expansions are well understood by many modelers, and this is an advantage not to be abandoned lightly. The other main advantage of using a consistent, non-adaptive basis system arises when a series of problems is being considered. The differences among corresponding analyses are then localized to the coefficients, instead of being partitioned out between the coefficients and the basis functions themselves.

The disadvantages arise in the case where the selected basis functions are not particularly well suited to the problem at hand. The Legendre basis, for example, is not a particularly

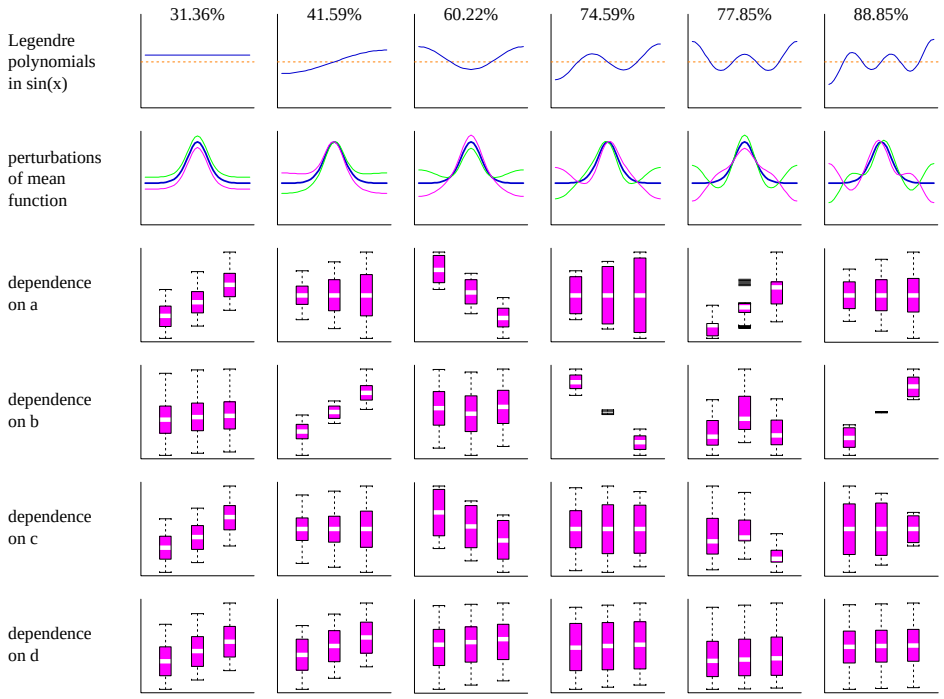

Figure 3. Dependence of the coefficients of the Legendre expansions for the four-parameter model on the parameters

Figure 4. Dependence of the coefficients of the Legendre expansions for the beam transport model on the parameters

good choice for a problem in which one of the main effects is neither symmetric nor antisymmetric, as for the 4-P example. The dispersion in the right-hand tail by comparison with the tight left-hand tail is not well captured by any single polynomial but spread out over several of them. The other disadvantage is that a relatively simple effect may be spread over several terms. For example, in this problem the effect of *b*, responsible for the left-right shift of the main peak, is spread out over all polynomials of odd order.

Figure 4 is a similar plot for the second through fifth Legendre polynomials for the beam transport example. (Because the uniform up and down shift accounts for less than one percent of the total variance, the polynomial of order zero is omitted from the plot.) As for the 4-P example, seven terms are needed to capture 90% of the total variance. Most of the action is in the even order, width-controlling terms, and the most important variable for these is *pyfac* (spread in y-momentum in the input beam), while *yfac* (spread in y-position in the input beam, not shown in Figure 4) is a distant second. The second important effect is left-right shift, which is controlled by *yshift* (the y-position shift of the input beam) and by *dy1* (the misalignment in the y-position of the first of the two quadrupole magnets.) Being carried along in Figure 4 for comparison with later methods are a couple of other scaling factors for the spread of input energy (*ptfac*) and phase shift in the input beam (*tfac*).

Adaptive bases computed by principal components analysis

The principal components of Y, considered as N observations in a T-dimensional space, are themselves T-vectors. They form an orthonormal basis for the T-dimensional space (or for a subspace of T-dimensional space, if N<T) that is specifically adapted to maximize the variance of the projection of the data onto the first basis vector, then onto the subspace spanned by the first and second basis vectors, etc. Thus expansions in the PC basis for sensitivity analysis should at least achieve some compression, avoiding one of the more serious problems with the Legendre polynomial, namely the allocation of a fairly simple effect (e.g., width changes or left-right shifts) to several components.

For the family of curves in Figure 1, the first principal component is basically an up-down shift, but unlike the first Legendre function this shift is not constant across all angles. (Refer to Figure 5.) The subspace spanned by this one function accounts for about 46% of the total variance in the family of curves, compared with about 31% for the Legendre polynomial of order zero. Like the zero-order Legendre coefficient, the coefficient of the first principal component depends on all four parameters. The second principal component for this example is a left-right shift accounting for another 34% of the total variance and controlled primarily by the *b* parameter. A similar amount of the total variance was spread across the Legendre polynomials of odd orders. The third principal component is devoted explicitly to the right-hand tail and accounts for 11% of the total variance. It is clearly controlled by the *d* parameter, something that could not be extracted from the Legendre analysis.

These first three terms capture over 90% of the total variance, compared to seven terms required by the Legendre analysis. The fourth component, which accounts for another 5% of the total variance, is a symmetric kurtosis or tail-fattening component depending most strongly on *a* and *c*.

For the beam profile example, just two terms account for over 90% of the variance (Figure 6). The first principal component is basically a widthing term, the second a left-right shift. From the point of view of the beam modelers, the third and fourth terms are also very interesting, both affecting the shapes of the tails. However, the box plots don't

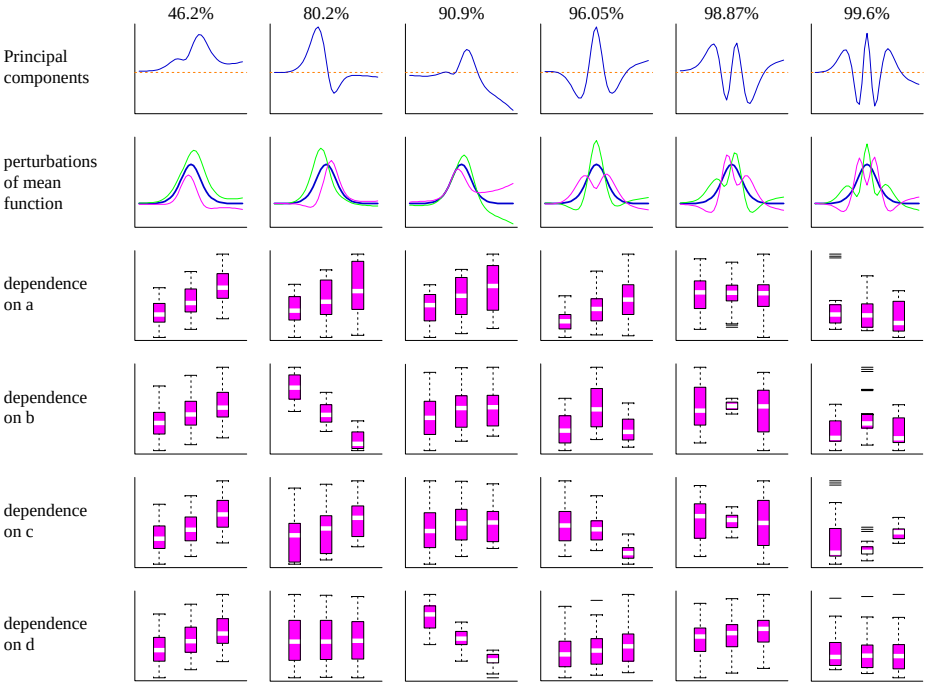

Figure 5. Dependence of the coefficients of PCA expansions for the four-parameter model on the parameters

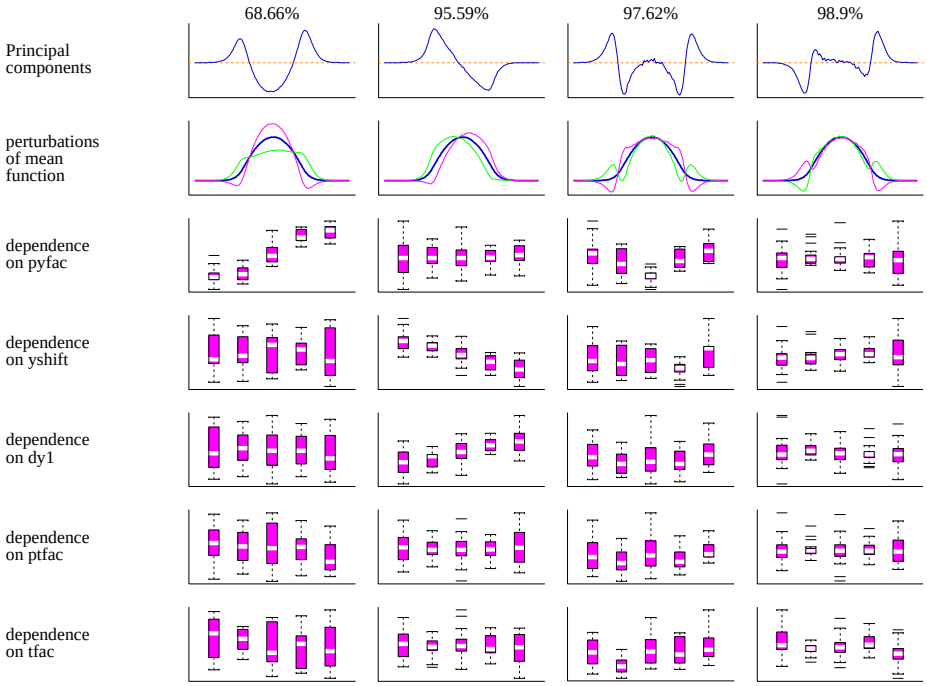

Figure 6. Dependence of the coefficients of PCA expansions for the beam transport model on the parameters

show exactly where these come from; the top variables for these components are the same as those for components 1 and 2, respectively.

Adaptive bases computed by partial least squares

Partial least squares (PLS) regression was invented to handle near-collinearity among the independent variables, which is not usually a problem in analyzing computer experiments, assuming a reasonable experimental design. Thus, PLS is really a technique for decomposing the design matrix. (For a review or PLS regression, see Frank and Friedman, 1993.) However, PLS simultaneously provides a transformation of the dependent variables in such a way that the first PLS component of the dependent variables is has the maximum variance that can be predicted by a linear combination of the independent variables. The second PLS component is computed using the residuals from the prediction of the first, and has the maximum variance that can be predicted by a second, orthogonal component of the independent variables, etc. So one might think of PLS as "peeking" at the explanatory variables while doing something that is similar to a PC analysis of the dependent variable. Note that while the PLS components of the independent variables are orthogonal, the PLS components of the dependent variables are not, in general.

While there is no *a priori* guarantee that PLS results will be interesting for functional sensitivity analysis, in the event they often seem to be fairly revealing. In particular, in the two examples they pull out some dependencies that were overshadowed by more important terms in both Legendre and principal component analyses.

For the four-parameter example, the PLS components (Figure 7) are somewhat more readily interpretable than the PCA components (Figure 5). The first component is an up-down shift of the middle of the curve, depending as before on all four parameters. (The first PLS component should be the same as the first PCA component if the independent variables are standardized, which is to be recommended; it is only with the extraction of the second component that the algorithms diverge.) The second PLS component is a left-right shift, almost entirely a function of *b*, compared to the second PCA component which had more substantial contributions from *a* and *c* as well. The third PLS component is pure right-hand tail, dependent on *d*. The fourth is primarily a widthing term, although it also includes a small left-right shift component, and depends on *a* and *c*. As there are only four input parameters, the PLS algorithm can provide only four component vectors, but this four-dimensional subspace captures almost 96% of the total variability in this family of curves, which is almost as much as the first four PCA components. By comparison, the first four Legendre components captured only about 75% of the total variance.

For the beam transport example (Figure 8), the first two PLS components are almost identical to the first two PCA components (Figure 6)—a widthing term and a left-right shift. The third component, no longer restricted to being orthogonal to the first two, turns out to be very similar to the first except for an arbitrary sign reversal, i.e., it is another widthing term. However, the explanatory direction for this component *is* required to be orthogonal to those already selected and turns out to have a strong dependence on *ptfac*, which controls the heterogeneity of the energy of the beam. Specifically, larger values of *ptfac* (increased heterogeneity or spread in the energy distribution) narrow the observed peak (the coefficients of this component are negative for larger values, corresponding to the direction of the red perturbation in the second line of the plot.) The fourth component is almost a pure tail-fattening component, and has a strong dependence on *tfac*, which controls the heterogeneity of the phase of the beam. Larger values fatten the tails. These

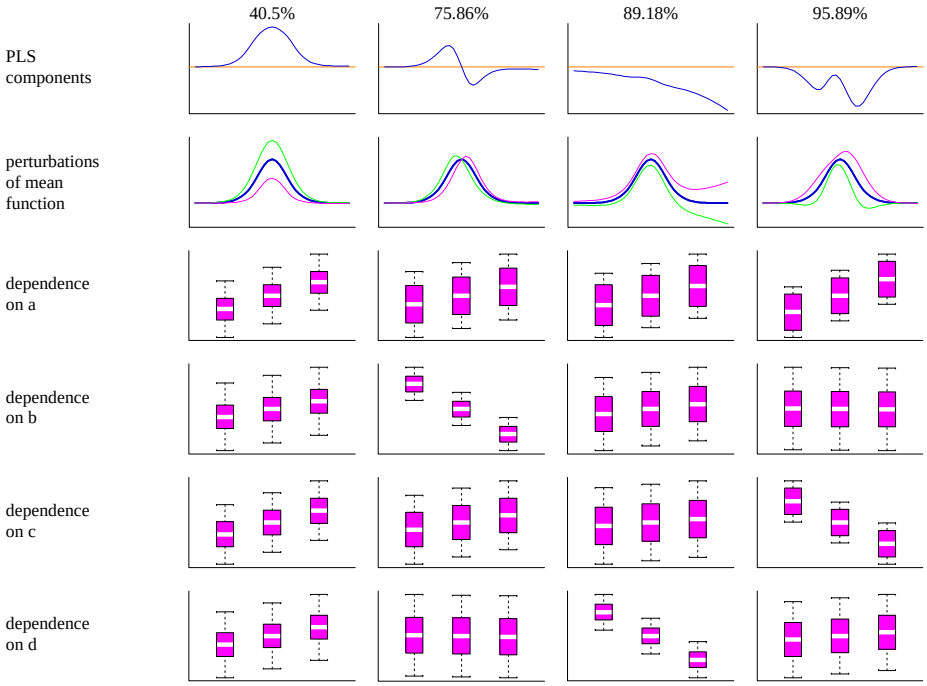

Figure 7. Dependence of the coefficients of PLS expansions for the four-parameter model on the parameters

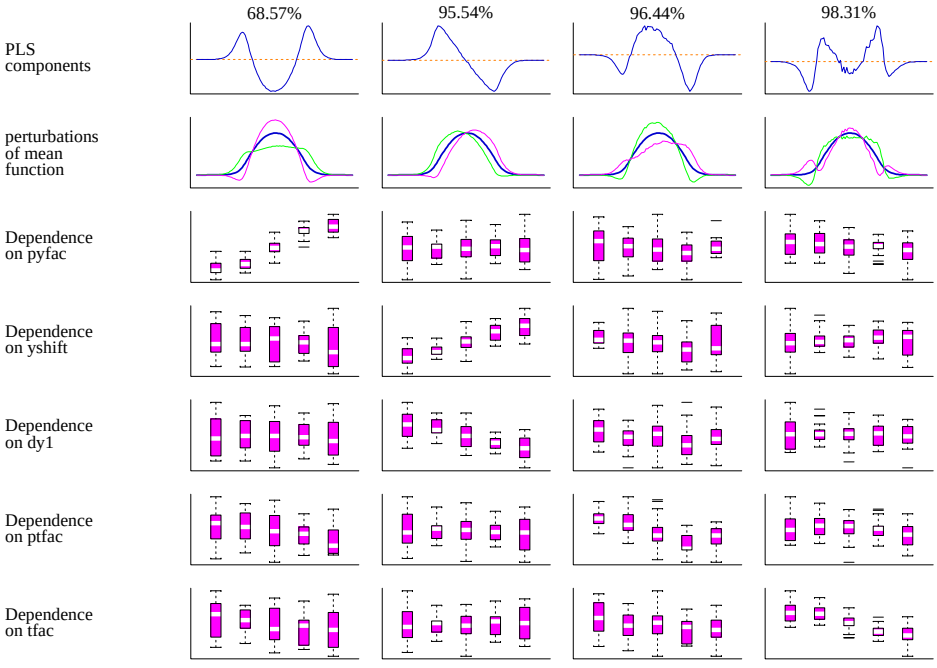

Figure 8. Dependence of the coefficients of PLS expansions for the beam transport model on the parameters

dependencies were not readily observable using the other functional transformations (although more sophisticated methods for sensitivity analysis than the simple box plots used here for illustration might have discovered them!)

The advantages and disadvantages of adaptive bases are pretty much the inverse of those for standard bases. The main advantage is good compression of the information; it is usually necessary to do sensitivity analysis on only the first few coefficients. The basis functions are also frequently more interpretable in physical terms. In a series of related problems, it may be interesting to study how the shapes of the component functions evolve (as well as their coefficients.) Of course, the down side to this is that shapes and coefficients are evolving simultaneously, which may lead to interpretation problems. In some cases it may make sense to pool all of the output functions for the series to extract a common set of principal components or PLS components, so that the evolution of their coefficients through the series can be studied in the same way as the evolution of the coefficients of a fixed basis set, such Legendre polynomials, could be examined.

Other considerations

Penalty methods can be used to enforce a degree of smoothness on adaptive basis functions. For example, some meaningless high-frequency information appears in the higher-order PCA and PLS components for the beam transport model output (Figures 6 and 8.) Orthonormality is lost when this is done, but the results are probably more interpretable, and curve comparison across problems, or between model output and noisy data, becomes easier. Ramsay and Silverman (1997) discuss the enforcement of smoothness in PCA (Chapter 7), and the technique is readily extended to PLS.

Curve registration may be needed or advisable when the parameters affect the time- or space-scale or when the functions not sampled at identical times in different runs. We may be interested in studying both the sensitivity of the scaling to the input parameters, independently of the variability in the functional outputs after adjusting for these scaling effects. Again, Ramsay and Silverman (1997) address this problem in detail, proposing a series of methods from parametric location and/or scale change, through feature or landmark registration methods, to the estimation of general monotonic transformation.

Summary

The purpose of this paper has been to suggest that sensitivity analysis for functional computer model outputs, correctly performed, is not significantly more difficult than for scalar outputs. The basic method is the expansion of the functional outputs in an appropriate functional coordinate system, i.e., in terms of an appropriate set of basis functions, followed by sensitivity analysis of the coefficients of the expansion using any standard method. The main art, then, is in choosing the appropriate coordinate system. We have considered both standard, pre-defined basis sets and data-adaptive basis sets. The examples tend to favor the latter because of the compression and interpretability of the results, but the former may have value, depending on the problem or set of problems and the customer.

References

Frank, I. E. and Friedman, J. H. (1993). A Statistical View of Some Chemometrics Regression Tools. *Technometrics* 35, pp. 109-135.

M. D. McKay, "Nonparametric Variance-based Methods of Assessing Uncertainty Importance," *Reliability Engineering and System Safety*, 57, 267-279, 1997.

Ramsay, J. O. and Silverman, B. W. (1997). *Functional Data Analysis*. Springer, New York